Math 164: Problem Set 10

Owen Jones

March 10, 2024

21.3 Let $f(x) = x_1^2 + x_2^2$, $g(x) = x_1^2 - x_2$, and $h(x) = x_1^2 + 2x_1x_2 + x_2^2 - 1$.

$$2x_1 + 2\lambda x_1 + 2\lambda x_2 + 2\mu x_1 = 0 \tag{1}$$
$$2x_2 + 2\lambda x_1 + 2\lambda x_2 - \mu = 0 \tag{2}$$
$$x_1^2 - x_2 \leq 0 \tag{3}$$
$$x_1^2 + 2x_1x_2 + x_2^2 - 1 = 0 \tag{4}$$
$$\mu \geq 0 \tag{5}$$
$$\mu(x_1^2 - x_2) = 0 \tag{6}$$

First, consider (3), (5), and (6). For (6) to hold, we need either $x_1^2 - x_2 = 0$ or $\mu = 0$.
Suppose $\mu = 0$. Using (1) and (2), we conclude $x_1 = x_2$. Plugging into (4), we obtain $[\frac{1}{2}, \frac{1}{2}]^\top$ and $[-\frac{1}{2}, -\frac{1}{2}]^\top$. However, $[-\frac{1}{2}, -\frac{1}{2}]^\top$ isn't a feasible point because $g([-\frac{1}{2}, -\frac{1}{2}]^\top) > 0$.
Thus, it suffices to show $y^\top L([\frac{1}{2}, \frac{1}{2}]^\top, -\frac{1}{2}, 0)y \geq 0$ for all $y \in \tilde{T}([\frac{1}{2}, \frac{1}{2}]^\top)$ $y \neq 0$. $\mu = 0$, so $\tilde{T}([\frac{1}{2}, \frac{1}{2}]^\top) = \{y : [1, 1]y = 0\} = \alpha[1, -1]^\top$.
$L([\frac{1}{2}, \frac{1}{2}]^\top, -\frac{1}{2}, 0) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$
$\alpha^2[1, -1] \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} [1, -1]^\top = 4\alpha^2 > 0$.
Suppose $\mu > 0$. Using (3), we conclude $x_2 = x_1^2$. Thus,

$$2x_1 + 2\lambda x_1 + 2\lambda x_1^2 + 2\mu x_1 = 0 \tag{7}$$
$$2x_1^2 + 2\lambda x_1 + 2\lambda x_1^2 - \mu = 0 \tag{8}$$
$$\left(x_1 + x_1^2\right)^2 = 1 \tag{9}$$

Subtracting (8) from (7) we get $2x_1 - 2x_1^2 + 2\mu x_1 + \mu = 0$ From (9) $2x_1 + 1 \neq 0$, so $\mu = \frac{2x_1^2 - 2x_1}{2x_1 + 1}$. From (9) $(x_1 + x_1^2) = \pm 1$, so $\mu = -2$ or $\mu = \frac{2 - 4x_1}{2x_1 + 1} \Rightarrow -\frac{1}{2} < x_1 < \frac{1}{2}$ which violate $\mu > 0$ or (9).
Thus, $[\frac{1}{2}, \frac{1}{2}]^\top$ is the only local minimizer.

21.12 The set of points satisfying KKT are those that satisfy

$$
\begin{aligned}
&x^\top Q + \mu^\top A = 0 \\
&\mu^\top (Ax - b) = 0 \\
&\mu \geq 0 \\
&Ax - b \leq 0
\end{aligned}
$$

From the second equation $\mu^\top Ax = \mu^\top b$. Postmultiplying the first equation by x gives us $x^\top Qx + \mu^\top Ax = 0 \Rightarrow x^\top Qx + \mu^\top b = 0$. However, $Q > 0$, $\mu \geq 0$ and $b \geq 0$, so $x = 0$ with μ s.t $\mu^\top b = 0$ is the only possible solution.

21.14

$$
\begin{aligned}
&c^\top + \mu^\top A = 0 \\
&Ax \leq 0 \\
&\mu \geq 0 \\
&\mu^\top Ax = 0
\end{aligned}
$$

Postmultiplying the first equation by x and subtracting the fourth equation gives us $c^\top x = 0$. Thus, $x = 0$ is a solution if there is one.

21.21 Suppose a solution exists. $f(x) = \frac{1}{2}\|x\|^2$, $h(x) = a^\top x - b$ $g(x) = -x$

$$
\begin{aligned}
&x + \lambda a - \mu = 0 \\
&\mu \geq 0 \\
&\mu^\top x = 0
\end{aligned}
$$

Premultiplying the first equation by $\mu^\top$ gives us $\mu^\top x + \lambda \mu^\top a - \|\mu\|^2 = \lambda \mu^\top a - \|\mu\|^2 = 0$. Also, premultiplying the first equation by $x^\top$ gives us $\|x\|^2 + \lambda x^\top a - x^\top \mu = \|x\|^2 + \lambda b = 0 \Rightarrow \lambda = -\frac{\|x\|^2}{b} < 0$ or $b = 0 \Rightarrow x = 0$. It follows $\frac{\|x\|^2}{b}\mu^\top a + \|\mu\|^2 = 0 \Rightarrow \mu = 0$ or else we have $\frac{\|x\|^2}{b}\mu^\top \geq 0$ by $a \geq 0$ and $\|\mu\|^2 > 0 \Rightarrow \frac{\|x\|^2}{b}\mu^\top a + \|\mu\|^2 > 0$.
Thus, $x = -\lambda a \Rightarrow x = \frac{b}{\|a\|^2}a$ to satisfy $a^\top x = b$. Hence, $x = \frac{b}{\|a\|^2}a$ if $a > 0$ or $x = 0$ if $a = 0$ is the unique solution.

21.25

$$
\begin{aligned}
&\mu \geq 0 \\
&\nabla f(x) + \mu \nabla g(x) = 0 \Rightarrow \nabla f(x) + \mu Dh(x)^\top h(x) = 0 \\
&\mu \|h(x)\|^2 = 0 \Rightarrow \mu \|h(x)\| = 0
\end{aligned}
$$

However, $h(x) = 0 \Rightarrow \nabla g(x) = 0$, so any feasible point x is not regular. Thus, KKT can't be used.

22.8 Yes. $f(x) = \frac{1}{2}\|Ax - b\|^2 = \frac{1}{2}x^\top(A^\top A)x - b^\top Ax + \frac{1}{2}\|b\|^2$ is quadratic. Since $F(x) = A^\top A \geq 0$, the function $f(x)$ is convex by theorem 22.5.
Next, we check the constraint set is convex. Pick feasible points x and y and $\lambda \in (0, 1)$. It follows $x_1 + x_2 + \cdots + x_n = 1$ and $y_1 + y_2 + \cdots + y_n = 1$. Thus, $\lambda(x_1 + x_2 + \cdots + x_n) + (1 - \lambda)(y_1 + y_2 + \cdots + y_n) = \lambda + (1 - \lambda) = 1$ and $\lambda x_i + (1 - \lambda)y_i \geq 0$, so $\lambda x + (1 - \lambda)y$ is a feasible point. Hence, the function is convex on a convex set.

22.12 (a)

$$x^\top Q - \lambda^\top A = 0$$
$$b - Ax = 0$$

Thus, $x = Q^{-1}A^\top\lambda$ and $b = AQ^{-1}A^\top\lambda$. Since, A is full rank, let $\lambda = (AQ^{-1}A^\top)^{-1}b$. Thus, $x = Q^{-1}A^\top(AQ^{-1}A^\top)^{-1}b$ is the only solution to the Lagrange condition.

(b) Yes, because $x^\top Qx$ is a positive definite matrix, $f(x)$ is convex, and because the constraint set is convex.

22.14 (a) Suppose x and y are feasible points in the constraint set and $\lambda \in (0, 1)$. It follows $\lambda a^\top x \geq \lambda b$ and $(1 - \lambda)a^\top y \geq (1 - \lambda)b$. Thus, $\lambda a^\top x + (1 - \lambda)a^\top y \geq \lambda b + (1 - \lambda)b = b$. Hence, the constraint set is convex.

(b) By KKT

$$2x - \mu a = 0$$
$$\mu \geq 0$$
$$\mu^\top(b - a^\top x) = 0$$
$$b - a^\top x \leq 0$$

$x \neq 0$ for equation 4 to hold because $b > 0$. This implies $\mu \neq 0$. Thus, for equation 3 to hold, we need $b - a^\top x = 0 \Rightarrow a^\top x = b$.

(c) By equation 1, we have $x = \frac{\mu}{2}a$. Then, $\frac{\mu}{2}a^\top a = b \Rightarrow \mu = \frac{2b}{\|a\|^2}$. Thus, $x = \frac{b}{\|a\|^2}a$ is the unique solution.

22.17 (a) $S_a := \{s : x_1 s^{(1)} + x_2 s^{(2)}, x_1, x_2 \in \mathbb{R}, s_i \geq a, i = 1 \cdots, n\}$
Let $\mathbf{a} = [a, a, \ldots, a]^\top \in \mathbb{R}^n$.
We can rewrite our problem in the form minimize $\frac{1}{2}(x_1^2 + x_2^2)$ subject to $x_1 s^{(1)} + x_2 s^{(2)} \geq \mathbf{a}$.

(b) By KKT

$$x_1 - \mu^\top s^{(1)} = 0$$
$$x_2 - \mu^\top s^{(2)} = 0$$
$$\mu \geq 0$$
$$\mathbf{a} - x_1 s^{(1)} - x_2 s^{(2)} \leq 0$$
$$\mu^\top(\mathbf{a} - x_1 s^{(1)} - x_2 s^{(2)}) = 0$$

22.18 (a) Want to show Ω is convex. Let u, v be probability vectors and $\lambda \in (0,1)$. It follows $u_1+u_2+\cdots+u_n = 1, u_i > 0$ and $v_1+v_2+\cdots+v_n = 1, v_i > 0$. $\lambda u + (1-\lambda)v = \lambda(u_1+u_2+\cdots+u_n) + (1-\lambda)(v_1+v_2+\cdots+v_n) = \lambda + (1-\lambda) = 1, \lambda u_i + (1-\lambda)v_i > \lambda 0 + (1-\lambda)0 = 0$. Hence, the constraint set is convex.

(b) Fix some p. We want to show $F(q) \geq 0$.

$$\frac{dp_i \log(\frac{p_i}{q_i})}{dq_i} = -p_i \frac{q_i}{p_i} \cdot \frac{p_i}{q_i^2} = -\frac{p_i}{q_i}$$
$$\Rightarrow \frac{d^2 p_i \log(\frac{p_i}{q_i})}{dq_i^2} = \frac{p_i}{q_i^2}$$
$$\Rightarrow F(q) = \begin{bmatrix} \frac{p_1}{q_1^2} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{p_n}{q_n^2} \end{bmatrix} > 0$$

because we have a diagonal matrix with all positive elements along the diagonal, so the eigenvalues of $F(q)$ are all greater than 0. Thus, $f(x)$ is convex on the constraint set.

(c) Fix some p. We obtain the convex minimization problem minimize $p_1 \log(\frac{p_1}{q_1}) + p_2 \log(\frac{p_2}{q_2}) + \cdots + p_n \log(\frac{p_n}{q_n})$ subject to $q_1+q_2+\cdots+q_n = 1, q_i > 0, i = 1\cdots, n$.

$$-\frac{p_i}{q_i} + \lambda = 0$$
$$q_1 + q_2 + \cdots + q_n = 1$$

From equation 1, $q_i = \frac{p_i}{\lambda}$. Using $p_1+p_2+\cdots+p_n = 1$, we obtain $\lambda = 1$. Thus, $q = p$ is the unique global minimizer. $f(p) = p_1 \log(\frac{p_1}{p_1}) + p_2 \log(\frac{p_2}{p_2}) + \cdots + p_n \log(\frac{p_n}{p_n}) = 0$ because $\log(1) = 0$. By definition of being a unique global minimizer, $f(q) > f(p)$ for all $q \neq p$. Hence, $D(q,p) \geq 0$, and $D(q,p) = 0$ iff $q = p$.